

1. What is meant by a non-manipulated characteristic of a participant? (Providing an example would help.) Then, tell me why such a variable would make causal claims difficult to make? [5 pts]

A non-manipulated characteristic of a participant is a quality that a participant brings with her or him to the experiment. Such characteristics would include gender, intelligence, motivation, self-esteem, age, height, education, socioeconomic status, etc. Use of such characteristics as factors is problematic because of the possibility that any result is due not to the characteristic, but to some other factor that is correlated with the characteristic. For example, suppose that you do a study that uses intelligence as a factor. Because intelligence does not occur in isolation, it may also be the case that more intelligent people have higher self-esteem, larger vocabularies, etc. Thus, if you observed differences between people with low intelligence and people with high intelligence, you wouldn't know for sure that the effect was due to intelligence or to some other variable associated with intelligence. As such, you are dealing with a third variable problem, even though the study is not a correlational design.

2. Professor Isabel C. Ewe has taught for several years and, based on her experience, believes that students who tend to sit toward the front of the classroom perform better on exams than students who tend to sit toward the back of the classroom. To test her hypothesis, she collects data from several of her classes, recording the student's final grade (A=4, B=3, etc.) and the typical row in which the student sat throughout the term (Row 1 = Front and Row 8 = Back). Thus, she conceives of her study as a single-factor independent groups design with two levels of the IV (Front and Back). She decides to analyze her data on StatView using ANOVA. Her output is shown below. Interpret the results of her study as completely as you can. [5 pts]

ANOVA Table for GPA

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Seat	1	1.225	1.225	13.667	.0007	13.667	.965
Residual	38	3.406	.090				

Means Table for GPA

Effect: Seat

	Count	Mean	Std. Dev.	Std. Err.
Back	20	3.090	.318	.071
Front	20	3.440	.280	.063

The first question was intended to be a hint for your answer to this question. Note that where a person sits is not manipulated by the experimenter, so it's a non-manipulated characteristic of the participant. Thus, you would not be able to conclude that the differences were due to the position of the student's seat. It could be that students with higher confidence sit up close, or students who don't hear as well. At any rate, you would be able to conclude that people who sit in the front of the class have significantly higher grades than people who sit in the back of class. (Note that no HSD is necessary, because there are only two groups.) What you couldn't conclude, however, is that the grades were due to the seating choice of the students.

3. Suppose that you are interested in determining if Motivation causes changes in performance on a problem-solving task. Your IV would be the level of motivation, with three levels: Low Motivation, Medium Motivation, and High Motivation. You decide to use the type of reward to produce the levels of motivation (\$.10 for each correct response in the Low Motivation group, \$.75 for each correct response in the Medium Motivation group, and \$1.00 for each correct response in the High Motivation group). You decide to use the number of problems solved within 30 minutes as your DV. Because of power considerations, you figure that you need at least 30 pieces of data for each condition. [20 pts]

a. Tell me what sort of design you would use and why.

Though I would be tempted to use a repeated measures design because of its power, I would probably go with an independent groups design to avoid any problems with using the three types of motivation on each participant. (Would the person get confused because he or she was getting paid a different amount on different trials? If the person were paid \$1 first, would they then really stop working if paid only \$.10?)

b. How many participants would you need?

For the independent groups design, I would need to run 90 participants. (n = 30)

[If you did choose to use a repeated measures design, you would need to run 30 participants. (Complete counterbalancing requires multiples of 6 participants.)]

c. Briefly, but explicitly, describe how you would conduct your study. (It may help to describe how you would run each participant.)

I would run the study in randomized replications. That is, I would randomly assign the first three participants to one of the three conditions (L, M, H motivation). I would use a different random order for the next three participants, etc.

Thus, as people entered the study, I would instruct them about the nature of the study, telling them that they would be paid (\$.10, \$.75, or \$1.00, depending on the condition) for each correct answer to the problems that would occur over the 30 minutes of the study. Then I would commence timing and give the participant the problems.

[If you did choose a repeated measures design, you would need to describe the 6 orders that you would use for your participants. Then, you would need to describe how you would instruct the participants about the various stages of the study.]

4. Dr. Richard Noggin is interested in the effect of different types of persuasive messages on a person's willingness to engage in socially conscious behaviors. To that end, he asks his participants to listen to each of four different types of messages (Fear Invoking, Appeal to Conscience, Guilt, and Information Laden). After listening to each message, the participant rates how effective the message was on a scale of 1-7 (1 = very ineffective and 7 = very effective). Complete the source table and analyze the data as completely as you can. [20 pts]

ANOVA Table for Message

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	20	27.000	1.350				
Category for Message	3	65.845	21.948	32.194	<.0001	96.583	1.000
Category for Message * Subject	60	40.905	.682				

Means Table for Message

Effect: Category for Message

	Count	Mean	Std. Dev.	Std. Err.
Fear	21	2.333	.796	.174
Conscience	21	3.000	1.000	.218
Guilt	21	4.714	1.102	.240
Information	21	2.952	.740	.161

With the significant overall effect, you would reject H_0 and conclude that the types of message differ in their impact on performance of socially conscious behaviors. To determine which messages actually differed would require computing HSD. In this case, with $q = 3.74$, $HSD = .67$. Thus, Guilt was significantly more effective than all other types of message. Appeals to Conscience were more effective than Fear messages (barely). Information laden messages did not differ from Fear messages or Appeals to Conscience. Of course, these are completely fabricated data, so don't take the results of this study to heart, but if the data were real, they would tell you that Fear-invoking messages were the most effective means of getting people to engage in socially conscious behaviors. (Let's hope that it's not true!)

But did you notice the problem with this study? (He's named Richard Noggin for a reason!) It's a repeated measures design with 4 levels. That means that you should use complete counterbalancing. Doing so would require 24 orders (and multiples of 24 participants), which means that he ran too few participants to have properly counterbalanced this study! (Even if he had used incomplete counterbalancing, 21 participants wouldn't work.)

5. First of all, define external validity and contrast it with internal validity (be very clear!!). What would you describe as the major point that Mook was making in his article on external invalidity? Mook argues that external validity may not be all that important when we are considering:

- whether something *can* happen, rather than whether it typically *does* happen;
- a prediction that something should happen in the lab, based on a theory;
- showing that a phenomenon is so powerful that it can happen even under unnatural conditions.

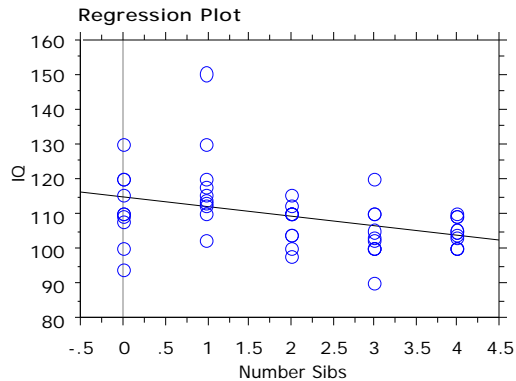
Using specific studies as examples, briefly describe the evidence from the article that illustrates *two* of the above conditions under which external invalidity may not be all that problematic. [10 pts]

Answers to this question should be fairly straightforward applications of the information from the Mook article.

6. Dr. Susan Mee is interested in the relationship between IQ and Number of Siblings. She is convinced that a "dilution of intelligence" takes place as siblings join a family (person with no siblings grew up interacting with two adult parents, person with one sibling grew up interacting with two adults+youngster, etc.), leading to a decrease in the IQ levels of children from increasingly larger families. She collects data from fifty 10-year-olds who have 0, 1, 2, 3, or 4 siblings and analyzes her data with StatView, producing the output seen below. Interpret the output as completely as you can and tell Dr. Mee what she can reasonably conclude, given her original hypothesis. What proportion of the variability in IQ is shared with Number of Siblings? If a person had 3 siblings, what would be your best prediction of that person's IQ? What about 5 siblings? On the basis of this study, would you encourage Dr. Mee to argue in print that Number of Siblings has a causal impact on IQ? Why or why not? [10 pts]

Regression Summary
IQ vs. Number Sibs

Count	50
Num. Missing	0
R	.406
R Squared	.165
Adjusted R Squared	.148
RMS Residual	9.281



ANOVA Table
IQ vs. Number Sibs

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	817.960	817.960	9.496	.0034
Residual	48	4134.760	86.141		
Total	49	4952.720			

Regression Coefficients
IQ vs. Number Sibs

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	114.880	2.273	114.880	50.532	<.0001
Number Sibs	-2.860	.928	-.406	-3.081	.0034

There is a significant negative linear relationship between Number of Siblings and IQ ($F(1,48) = 9.496$, $p < .05$). Thus, the results are consistent with Dr. Mee's hypothesis. With $r^2 = .165$, you would conclude that the two variables share about 17% of their variability. Thus, a good deal of the variability in IQ is attributable to other variables. If a person had 3 siblings, you would predict $IQ = 106.3$ (using the regression equation). If a person had 5 siblings, you couldn't really predict, because you didn't observe any people with that many siblings. And, of course, you would advise Dr. Mee to avoid making a causal claim based on this correlational study! The causal factor could be a third variable (e.g., larger families may have less money, which would provide fewer educational resources for each child).